

Segmentation of nuclei and cells using membrane related protein markers

C. ORTIZ DE SOLORZANO, R. MALLADI, S. A. LELIÈVRE & S. J. LOCKETT

Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, U.S.A.

Key words. Cell surface, confocal, geometrical flow, image segmentation, integrin, lamin, nuclear lamina, partial differential equation (PDE).

Summary

Segmenting individual cell nuclei from microscope images normally involves volume labelling of the nuclei with a DNA stain. However, this method often fails when the nuclei are tightly clustered in the tissue, because there is little evidence from the images on where the borders of the nuclei are. In this paper we present a method which solves this limitation and furthermore enables segmentation of whole cells. Instead of using volume stains, we used stains that specifically label the surface of nuclei or cells: lamins for the nuclear envelope and alpha-6 or beta-1 integrins for the cellular surface. The segmentation is performed by identifying unique seeds for each nucleus/cell and expanding the boundaries of the seeds until they reach the limits of the nucleus/cell, as delimited by the lamin or integrin staining, using gradient-curvature flow techniques. We tested the algorithm using computer-generated objects to evaluate its robustness against noise and applied it to cells in culture and to tissue specimens. In all the cases that we present the algorithm gave accurate results.

1. Introduction

Computer-based segmentation and reconstruction of cell nuclei from microscope images allows the study of individual cell nuclei within their natural tissue context. This is vital for understanding the molecular basis behind how tissue develops, how the cellular organization of tissue remains stable through the adult life of the organism, and what goes wrong in cancer development, which is associated with a loss of tissue organization.

Traditionally, nuclear segmentation has been done using images of nuclei counterstained with fluorescent DNA probes (e.g. PI, DAPI, Hoechst) which stoichiometrically

bind to the DNA, producing images with high contrast between nuclear areas and surrounding background. Manual methods, based on drawing with a mouse the boundaries of the nuclei, are the gold standard, but the amount of interaction required is too high for most applications where many nuclei must be segmented. A number of algorithms have been developed to automatically or semi-automatically segment nuclei in 2D (Lockett & Herman, 1994; Dow *et al.*, 1996; Malpica *et al.*, 1997) and 3D images (Rigaut *et al.*, 1991; Irinopoulou *et al.*, 1997; Rodenacker *et al.*, 1997; Lockett *et al.*, 1998; Ortiz de Solorzano *et al.*, 1999). With more or less sophistication, all these algorithms rely on the existence of unstained areas (background) between the nuclei. As a result, these algorithms perform particularly well in specimens that contain isolated or not too clustered, morphologically normal nuclei, but fail otherwise. Unfortunately, real tissue specimens and many cell cultures seldom fulfil these requirements at the resolution of optical microscopy.

We have explored an alternative avenue to DNA stain based nuclear segmentation, which uses stained proteins located on the nuclear lamina, a protein mesh that underlies the inner membrane of the nuclear envelope. Therefore, using fluorochrome-tagged antibodies against one or a combination of lamins, a detectable signal between nearest neighbour nuclei can be obtained in 2D or 3D (confocal) images of cell cultures or tissue. Thus, the domain occupied by each nucleus is nicely delimited, no matter how clustered the nuclei might be (see Figs 4a and c).

We also want to segment whole cells in order to quantify the expression of cytoplasmically localized proteins and mRNAs. We propose to do this using cell surface-bound protein markers. This is an unexplored area that extends the method for segmenting nuclei proposed above. In this case, we believe that the use of cell surface markers is the only approach because cells always touch each other in solid tissue and are thus 'inherently clustered'. As cell surface

Correspondence: Dr Stephen J. Lockett, SAIC-Frederick, PO Box B, Frederick, MD 21702, U.S.A. Tel.: +1 301 846 5515; fax: +1 301 846 6552; e-mail: slockett@ncifcrf.gov

markers we used integrins, which are a family of cell-surface transmembrane proteins that facilitate cell-cell and cell-extracellular matrix interaction (Bosman, 1993). Specifically, we used either alpha-6 or beta-1 integrins, because of their ubiquitous expression: they have been localized on almost all epithelial cell types as well as the endothelial cells of capillaries and in mast cells (Koukoulis *et al.*, 1991).

To the best of our knowledge, there is only one previous report on segmentation based on cell surface markers (Dow *et al.*, 1996). They used multiple fluorescent staining (DNA + three monoclonal antibodies against surface antigens: CD3, CD8 and HLA-DR) to identify tumour infiltrates in melanoma tissue sections. They first detected cell nuclei using a watershed-based algorithm applied to the image of DNA staining and then used the segmented nuclei as seeds for a snake spline adjustment (Kass *et al.*, 1988) directed to the labelled surfaces of the cells. What distinguishes our work from theirs is that we want to do segmentation of both nuclei and whole cells using surface markers when the standard DNA stain-based algorithms do not provide accurate results. Furthermore, we wanted to avoid as much as possible the limitations of the snake approach, which requires a close initialization (proximity between the initial surface and the surface to reach) and assumes surface smoothness.

The approach we have taken here for segmentation of surface labelled specimens is in part inspired by methods used in the medical imaging field to solve 'cavity detection' problems (e.g. Balzer *et al.*, 1998; Dastidar *et al.*, 1999; Kikinis *et al.*, 1999). Some of these problems were addressed using region growing algorithms, which assume an internal homogeneity and may require a priori knowledge about the shape of the objects. In our images, the noise level, the inherent heterogeneity of the staining and the variability of nuclear shapes and sizes dissuaded us from using this approach. An alternative method is edge finding algorithms that use gradient information to detect discontinuous edges that are subsequently joined using morphological matching or energy optimization techniques (Cohen & Cohen, 1993) to find the surface that best matches the gradient map. However, these methods assume local surface smoothness and require a precise initialization. Therefore we did not use them.

Encouraged by recent advances on the application of partial differential equations (PDEs) to image filtering and segmentation (Alvarez *et al.*, 1993; Romeny, 1994; Malladi *et al.*, 1995; Malladi & Sethian, 1996, 1997, 1998; Mikula *et al.*, 1997; Elmoataz *et al.*, 1998; Sochen *et al.*, 1998; Sarti *et al.*, 1999) using gradient-curvature driven flows, we have also applied them to our problem.

In this paper we present and demonstrate our segmentation approach, which is based on gradient-curvature driven flows. It first requires finding an internal seed that uniquely identifies each cell/nucleus in the image. Then each seed's

surface is dilated until its surface aligns with the nuclear lamina or cell membrane surface. By contrast with the energy optimization methods, this method does not have strict initialization requirements in that the size, shape and position of the seed are not as critical in providing an accurate segmentation. Furthermore, the flows can be adapted to local image characteristics, such as curvature, edge strength and direction, etc. Advantages from the implementation point of view are that a discrete approach can be used to approximate the solution of the flow equation that describes the movement of the surface, and that geometrical parameters such as the curvature of the surface can be easily extracted from the higher dimensional function.

2. Materials and methods

2.1. Sample preparation

2.1.1. Nuclear lamina staining. HMT-3522 S1-50 human immortalized mammary epithelial cells (HMECs) were propagated as monolayers in chamber slides in chemically defined medium (Petersen *et al.*, 1992). After 5 days, the culture showed multiple islands of clustered cells. Cells were fixed either in methanol-acetone (1 : 1) at -20°C for 10 min or in paraformaldehyde (following 0.5% Triton permeabilization) for 30 min at 4°C . Immunostaining was performed using a protocol already described (Lelièvre *et al.*, 1998). Mouse IgG1 anti-laminB antibody was purchased from Matritech (Cambridge, MA) and secondary anti-mouse IgG1 FITC-conjugated antibody was purchased from Jackson Laboratories (West Grove, PA, U.S.A.).

2.2.1. Cell surface staining. The same 2D cell culture described for the nuclear staining was also used for cell surface staining with antibodies against either alpha-6 integrin or beta-1 integrin from Pharmingen (San Diego, U.S.A.).

Finally, tissue sections were prepared as follows: Nulliparous mouse mammary gland was fresh frozen in embedding medium and $5\text{ }\mu\text{m}$ cross-sections were fixed in methanol/acetone (1 : 3) for 5 min prior to incubation with the beta-1 integrin primary antibody (Pharmingen) as previously reported (Barcellos-Hoff, 1993).

The integrin antibodies for both cultured cells and tissue were conjugated with FITC.

2.2. Image acquisition

Confocal images from lamin and integrin-stained specimens were acquired using a laser scanning confocal microscope 410 (Carl Zeiss Inc., Thornwood, NY) equipped with an Axiovert 100 inverted microscope (Zeiss), a $63\times$, 1.4 NA plan-Apochromat objective lens (Zeiss) and an Argon/Krypton (Ar/Kr) laser. The specimens were excited using

the 488 nm line of the Ar/Kr laser. The resolution of the images was close to the theoretical maximum (0.2 μm in the lateral (x, y) direction, 0.5 μm in the axial (z) direction). The images had an average of 40 z -sections and were manually selected to contain clusters of cells. Images were stored in the ICS image format (Dean *et al.*, 1990) and transferred to a Unix workstation for archiving and analysis.

2D images from integrin-stained tissue sections were imaged on an Axioplan Microscope (Carl Zeiss Inc.), using a 40 \times , 0.75 NA Plan_Neofluar objective lens (Zeiss), and a 100 W halogen lamp (Osram, Germany). Images were stored in TIFF format.

2.3. Image analysis (nuclear segmentation)

The following sections describe how the algorithm works in 2D.

2.3.1. Locating a seed inside each nucleus. The objective is to find a single seed inside each nucleus, which is any closed area of the image entirely within the boundaries of its nucleus. Finding more than one seed inside a nucleus causes oversegmentation (the nucleus will be divided into more than one object). If no seed is found within a given nucleus, the algorithm will not be able to detect it.

Our seed-finding method is inspired by the Hough transform (HT) algorithm (Ballard, 1981), and a full description of the method can be found in Ortiz de Solorzano *et al.* (1999). Briefly, the method is as follows. First, the magnitude and direction of the gradient at each pixel in the original image is calculated using a 3×3 Sobel filter, producing the image I_G . Then a new image, I_R , which is initially empty, is generated as follows. For each pixel in I_G its intensity is added to the intensities of the two pixels in I_R that are a distance R away in the positive and negative directions of the gradient. This procedure is repeated for a range of values of R , resulting in a set of images I_R . The rationale for this approach is that for quasi circular objects (see Fig. 1a) the intensities from I_G will accumulate near the centre of the circle in image I_R when R is approximately the radius of the circle (see Fig. 1b). By contrast, image I_R will show a diffuse pattern when R does not approximate the radius. In other words, the image I_R that was generated from the correct value for $R(I_{Rest})$ is the one that is most 'in focus'. This image can be automatically identified from the set of images I_R by use of a focus measurement algorithm. For each image, I_R this algorithm calculates the summation

of the magnitude of the gradient at every pixel and reports the images with the highest sum as being the most 'in focus'. Here, gradient was calculated as the difference in the intensities of adjacent pixels in the x direction.

The clusters of signals in image I_{Rest} (Fig. 1b) are used as seeds for initialization of the boundary finding process described in the next section. This image is then Gaussian filtered with $\sigma = R/4$ to remove noise (Fig. 1c) and to take into account that not all objects are the same size or perfectly circular. Then the peaks of the image are extracted using a dome extraction algorithm which uses morphological reconstruction (Vincent, 1993) (Fig. 1d). Each peak is considered a seed, and its surface is dilated as described below to the true surface of the object.

Alternatively, seeds can also be interactively defined by a mouse click somewhere in the object. In our experience the location of the seed in the object was not critical.

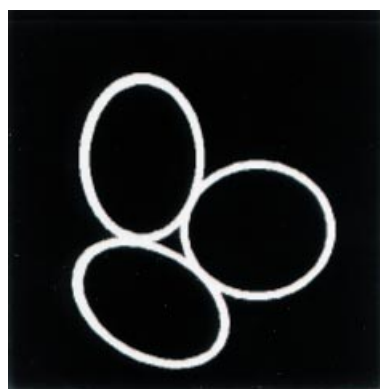
2.3.2. PDE based segmentation of nuclei and cells. Our segmentation method consists of moving the surface of each seed until it adopts the form of the enclosing nucleus or cell, as delineated by the surface staining. This is accomplished by enunciating and solving the equation that describes the movement of the surface from the initial position, defined by the seed finding algorithm above, to its final position, where the seed's surface aligns with the surface. We allow the surface to move only in the normal direction at a speed that depends on a force term $F(x)$ which is tuned to local characteristics of the surface along with properties of the original image. The particular details of the force term are described below.

To solve the equation of the movement of the surface under the force term $F(x)$, we adopted the level set approach, where the moving surface $\gamma(x)$ is embedded as the zero level set of a higher dimensional function $\psi(x)$ (Sethian, 1997). Then the equation of the movement of the function can be written and solved as a PDE:

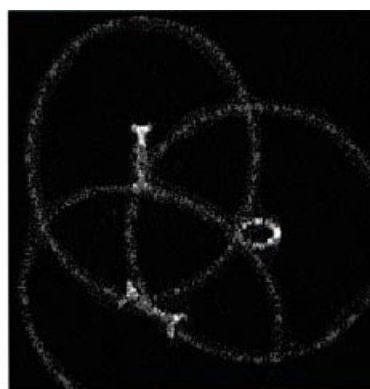
$$\psi_t + F|\nabla\psi| = 0 \quad (1)$$

with a given initial condition $\psi(x, t = 0) = \psi_0$ which we chose to be the distance transform $d(x)$ of the seed's surface $\gamma(x)$, assigning negative distances for pixels inside the evolving object and positive distances for pixels outside it. The level zero curve $\psi(0, t = T)$ holds the position of the evolving seed's surface at any time point T in the evolution of the function. The rest of the level sets are iso-distance curves $\psi(d, t = T)$, which contain all pixels located a distance d from the level set 0. We can see the evolution of

Fig. 1. Example of the use of the algorithm using a computer-generated image (a) simulation of a cluster of three surface-stained nuclei. Seeds are detected using a method based on the Hough transform (HT), where gradients of the original image are projected using an estimate of the object size (b). The HT image (b) is then smoothed using a Gaussian filter (c) and then seeds are extracted from (c) using a morphological peak detection algorithm (d). Each seed, which uniquely identifies one of the nuclei (e), is then moved using the sequence of flows: initial expansion (f), free expansion (g) and surface wrapping (h).



(a)



(b)



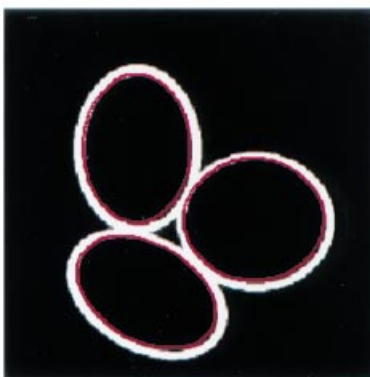
(c)



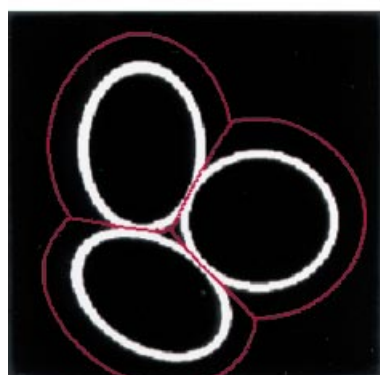
(d)



(e)



(f)



(g)



(h)

the function $\psi(x)$ as the evolution of a front of iso-distance surface. Therefore we will use the term front to refer to the function $\psi(x)$.

The equation of the movement of $\psi(x)$ is solved using finite differences and an iterative scheme. Moving the front m steps is equivalent to calculating $\psi(d, t = m\Delta t)$ for all the pixels in the image. The evolution of the seed's surface is defined by those pixels (i, j) in which $\psi_{ij}(x) = 0$ after m iterations of the algorithm, being Δt the unit (quantum) movement per iteration. Δt is an important parameter, since it determines the speed of the movement. If Δt is too large the front cannot converge to a stable solution or voids the effect of the force term. If Δt is too small, errors do not arise, but the front evolves very slowly, with the consequent computational and time cost. Our approach to select Δt is empirical. However, once a value has been found appropriate for a given image, it can then be used for similar images. For more details about the implementation see (Osher & Sethian, 1988; Sethian, 1997).

To reduce computation time, we used a narrow band method (Sethian, 1997) that consists of updating the front only for those pixels located within a distance d_{\max} of the level set 0. This approach substantially reduces computation time by not spending time in areas which are far away from the zero level set. The distance transform has to be rebuilt after several narrow band iterations. In our case we rebuild it every 10 iterations, since we found that number a good trade-off between computation time and accuracy.

If an image contains n nuclei, each seed is embedded in an independent function ψ^i , ($1 \leq i \leq n$) which is moved independently of the other fronts. In order to prevent the fronts from crossing when they are expanding, their movement is limited by the position of the other fronts. Accordingly, every time a front is moved one step, the outcome is considered as a trial function. The final movement of the front is computed as the maximum between the value of the trial function and the value of the other fronts.

$$\psi_{m+1}^i = \max \{ \psi_{m+1}^{i(trial)}, \psi_m^j \}, 1 \leq j \leq n, i \neq j \quad (2)$$

2.3.3. Sequence of flows. To solve our particular problem, we devised a sequence of flows to accomplish the goal. What follows is a description of the sequence of flows, their effect being also illustrated on a synthetic image (see Fig. 1) which resembles the type of image that we want to segment. Figure 1(d) shows the initial seeds, which correspond to the zero level sets of ψ_0^i , ($1 \leq i \leq 3$) at time zero.

Flow 1. Initial expansion. The initial flow moves the front towards the internal surface. It follows a particularization of the general Eq. (1):

$$\psi_t - g \cdot (1 - \varepsilon H) \cdot |\nabla \psi| - \beta \nabla g \cdot \nabla \psi = 0 \quad (3)$$

The first term, g , attracts the surface towards areas of the image with high gradient. This is done through the nonlinear gradient function g :

$$g(x) = e^{-\alpha |\nabla(G * I_0(x))|}, \alpha > 0 \quad (4)$$

where $G * I_0(x)$ is the convolution of the original image $I_0(x)$, with a Gaussian function (G). $x \equiv (x_1, x_2)$ for a two dimensional image. The standard deviation of G , σ determines the size of the smallest object preserved by the filter. The effect of g is to speed up the flow in those areas where the image gradient is low ($g \equiv 1$) and slowing it down where there is high gradient ($g \equiv 0$). The parameter α determines the sensitivity of the flow to the gradient.

The effect of g is modulated by the second term $(1 - \varepsilon H)$, which contains an inflationary term $(+1)$, enhanced or opposed by a curvature term (εH) that regularizes the surface by accelerating the movement of those parts of the surface that are behind the average of the front (and therefore have a negative curvature) and slowing down the advanced parts of the flow (parts with positive curvature). The parameter ε determines the strength of the regularization: a low ε allows creating or maintaining sharp corners on the fronts, while a high value will smooth out front irregularities and in the extreme case will result in only circular objects. In practice an intermediate value of ε was chosen so that concavities in nuclear and cell borders, which have low curvature, are followed, but bright spots from punctate staining, small gaps in the staining and noise are smoothed over.

The third term in Eq. (3), $\beta \nabla g$ is a parabolic term that enhances the edge effect, once an edge is reached due to the effect of the first two terms. It aligns all the level sets with the ideal gradient, which would be a perfect step function, centred at the point of maximum gradient in the original image.

Because of the gradient, the front slows down almost to a stop when it reaches the inner face of the nuclear surface. This can be used as the stop condition for the flow when the algorithm is run in an unsupervised way. It can be implemented by checking the volume increase after each iteration, and setting a minimum threshold of volume change that when reached will interrupt the flow. Otherwise, a conservatively high number of iterations can be done that will ensure that the front will reach the internal side of the stained surface. The result of this initial flow can be seen in Fig. 1(f).

The flow described above produces an initial, but underestimate of the nuclear shapes, but does make subsequent flows independent from the size or shape of the initial seeds. The purpose of the subsequent flows is to correct for this underestimation.

The values of the parameters used in this flow, α , β , ε , Δt , are determined empirically. In our experience, a given set of parameters can be used for images within a broad range of

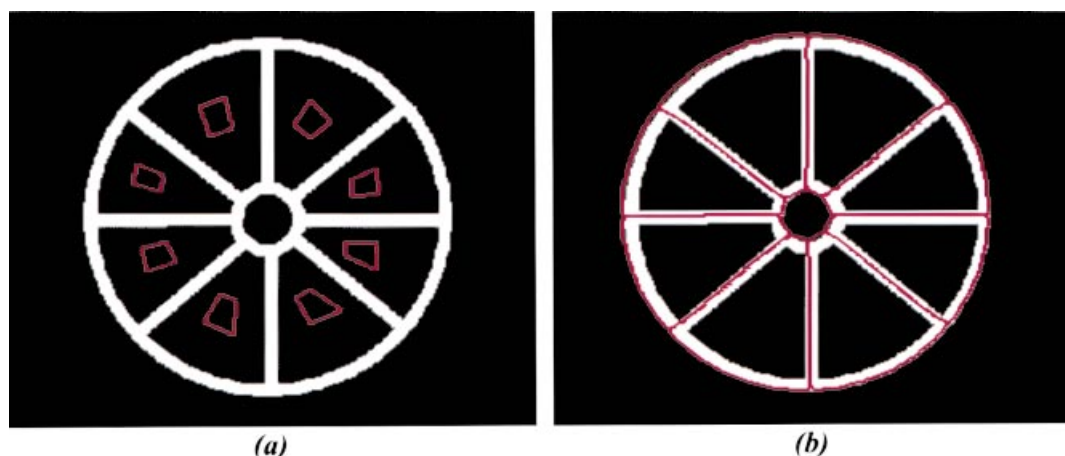


Fig. 2. Computer-generated object simulating a group of integrin-stained epithelial cells. (a) Manual initialization of the segmentation. (b) Final segmentation result after applying the algorithm described in the text.

image characteristics (gradient, noise level, etc.), as will be shown below. Variations of the parameters or the image properties will mainly alter the speed of the segmentation (how fast the front converges to the final position) but not the accuracy of it.

Flow 2. Free expansion. The second flow detaches the front from the inner surface of the staining and allows it to expand freely and independently from the gradient in the original image. The equation that describes the movement is:

$$\psi_t + |\nabla\psi| = 0 \quad (5)$$

The only limit to the expansion of a front is the position of the other fronts. The expansion is allowed only for a number of steps that ensure that all the flows move beyond the outer surface. The number of steps can be again empirically calculated based on the membrane width or a fixed value can be used that will ensure the desired evolution of the front. The result can be seen in Fig. 1(g).

Flow 3. Surface wrapping. Finally, the surface must be moved inwards until it finds the external stained surface. To do this we use the flow:

$$\psi_t - g(-1 - \varepsilon H)|\nabla\psi| = 0 \quad (6)$$

which is similar to Eq. (3), but with a negative advection value, which moves the front inwards, and with the function g dependent on the intensity alone of the image, and not on its gradient.

$$g(x) = e^{-\alpha|G^*I_0(x)|}, \alpha > 0 \quad (7)$$

The last term in Eq. (3) is also removed. The effect of this flow is an inward movement opposed by high intensity values, as those represented by the maximum intensity of the stained lamina. The result is shown in Fig. 1(h).

The characteristics of the staining of the cellular surface based on integrins are not very different from those of the nuclear staining using lamin antibodies. Therefore, we used the same sequence of flows, with identical parameters as the ones used for the image in Fig. 1. Figure 2 shows a computer model which represents a group of cells forming a structure that resembles a duct in mammary tissue. Figure 2(a) is the original image with the seeds interactively drawn. Figure 2(b) shows the results after using the same sequence of flows as used for Fig. 1.

2.4. Testing

2.4.1. Robustness against noise and changes in gradient. First we checked the accuracy of the algorithm and its robustness against noise and gradient variations. In the test, we used the same computer-generated objects used to illustrate the way the algorithm works (Fig. 1). The image was segmented multiple times following changes in surface intensity and under different levels of Gaussian noise. Following the method described in Ortiz de Solorzano *et al.* (1999), we tested the accuracy of the algorithm by measuring the average distance of the true surface to the surface obtained using our segmentation algorithm.

2.4.2. Resistance to surface discontinuities. We evaluated the resistance of the segmentation against discontinuities in the surface staining. This was done using the computer-generated object shown in Fig. 3, by introducing gaps of increasing size (3, 6 and 9 pixels), and determining if the flow converged to the desired surface or escaped through the holes. Given that ε is the parameter that controls the local curvature of the flow, we determined the minimum value of ε that prevented the flow from escaping through

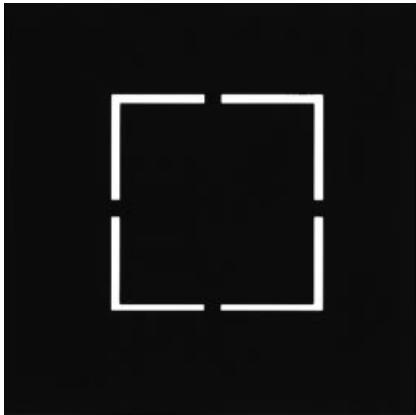


Fig. 3. Computer-generated object used to determine the robustness of the algorithm against surface discontinuities. The gaps in this image are 6 pixels wide.

the holes. The values of the rest of the parameters were the same as those used above.

2.4.3. Testing with real images. The algorithm was tested on real images of lamin- and integrin-stained cells in culture and tissue to demonstrate the practical application of the approach. The images were selected to cover a range of situations of staining quality and image noise which gave us confidence that the algorithm would work on images from a wide variety of specimens.

3. Results

Unless specified, the values for the parameters used in all the experiments (see Table 1) were the same as those used for the segmentation of the computer-generated image shown in Fig. 1.

3.1. Robustness against noise

Table 2 shows the average distance in pixels between the true surface of the computer-generated, noise-free object shown in Fig. 1(A), and surfaces detected by the algorithm

Table 1. Parameters used in the flows. Gradient attraction (α), edge effect (β), curvature-limiting effect (ε), step size of the numerical approximation to the solution of the PDE flow (Δt), number of iterations of the approximation (# iter).

	α	β	ε	Δt	# iter
Flow 1	0.1	0.2	0.005	0.0005	200
Flow 2	0	0	0.005	0.0005	20
Flow 3	0.05	0	0.005	0.0005	40

Table 2. Robustness of the algorithm against noise. The table shows the average surface error after segmenting the image in Fig. 1 under different noise conditions. Six noise levels (1–6) were considered, by adding Gaussian noise with 0 mean and increasing standard deviation (SD). The SD values for the signal and background are shown in rows 2 and 3, respectively. Decreasing values of image quality were also considered, by reducing the gradient between background and signal from 250 to 100 in steps of 50 intensity values. The error values are shown in rows 4 to 7.

Image	1	2	3	4	5	6
Signal SD	34.21	35.34	39.87	43.16	41.17	49.19
Back SD	12.63	17.4	22.49	25.27	27.8	31.6
Contrast: 250	0.3	0.45	0.57	0.69	0.84	0.98
Contrast: 200	0.2	0.37	0.53	0.69	0.89	1.01
Contrast: 150	0.2	0.35	0.47	0.8	0.92	1.01
Contrast: 100	1.13	1.56	1.32	1.2	1.65	1.56

of the same object after varying the contrast and addition of noise. Contrast was defined as the difference in grey level between the intensity on the surface and background intensity. As shown in Table 2, the average distance remained below one pixel for contrasts of 150 or greater, except for the highest noise levels tested (image 6). For all other cases, the average distance remained under two pixels.

For comparison of the values shown in Table 2 to those for images of real tissue, the noise and gradient values of our real images (measured from many images) were always in the range of noise levels for images 1 and 2 in Table 2, with gradient values higher than 150. Therefore we can expect an average distance (accuracy) lower than 0.37. Only very dim images with poor noise conditions will generate higher errors, but the use of analysing those images is at least dubious, given their poor image properties.

3.2. Resistance to surface discontinuities

Table 3 shows the minimum value of ε required to prevent the flow from escaping through holes of increasing size in the ‘surface’ represented by the image in Fig. 3.

Table 3. Robustness of the algorithm against surface discontinuities. The table shows the minimum value of ε that was used to prevent Flow 1 from escaping through the holes inserted in the image in Fig. 3.

Gap size	3 pixels	6 pixels	9 pixels
Parameter ε	0	0.005	0.01

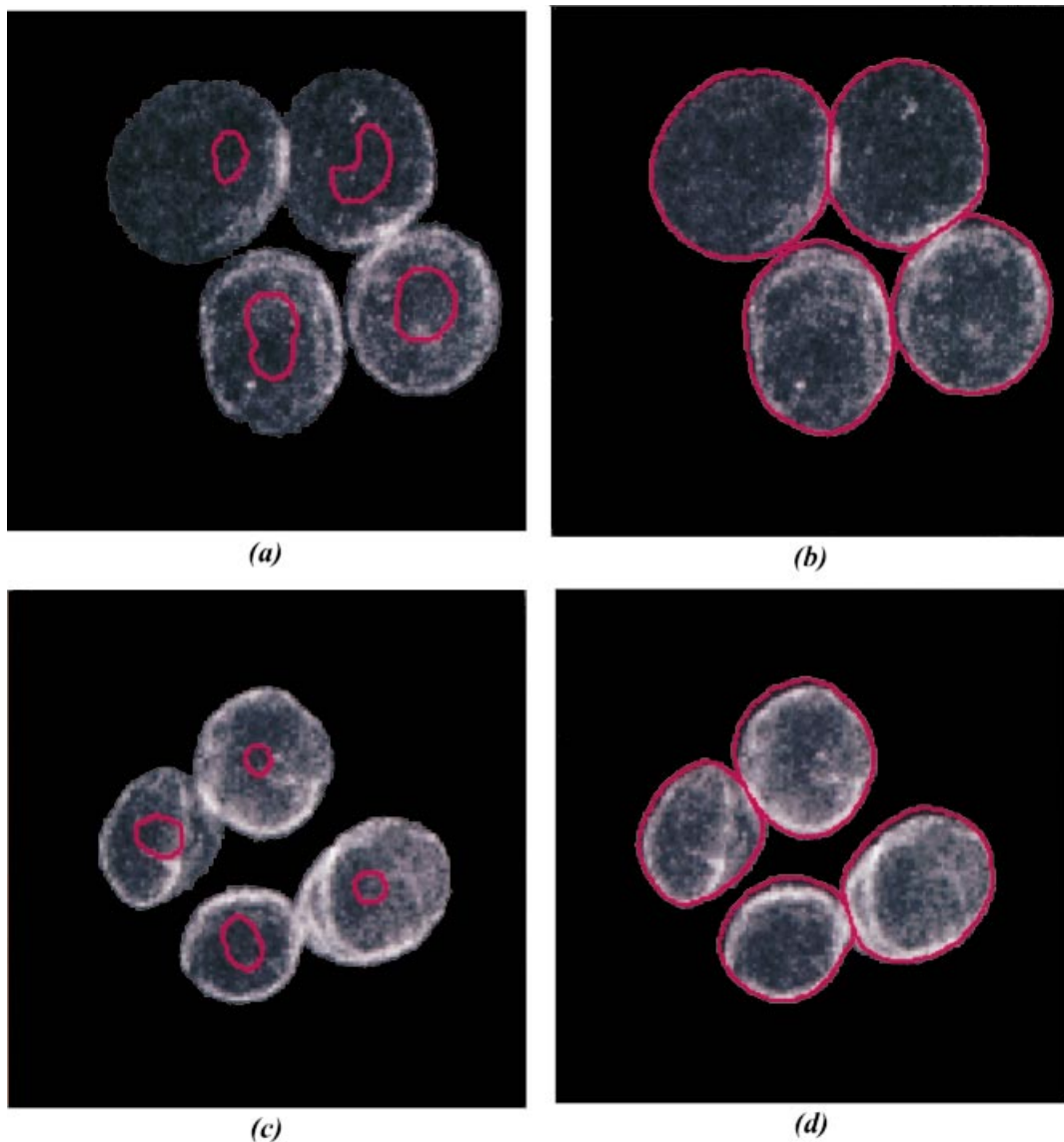


Fig. 4. Examples of the segmentation results on real images from cultured lamin-stained cells. (a) and (c) show the original images with the automatically detected seeds. (b) and (d) show the results, superimposed on the original images.

As we can see, for the hole sizes studied, we always found a value of ε that prevented the flow from flowing through the holes. An obvious consequence of increasing the value of ε is that the flow will follow the surfaces less accurately if they are highly convoluted. This poses a trade-off choice between surface accuracy and 'hole' filling of the flow that should be solved by the user. We also point out that the watershed algorithm will fail to accurately find these surfaces at the location of holes (unless the hole is exactly halfway between the two opposing starting points).

3.3. Experiments with real images

So far, the lamin staining in biological specimens has been observed to be high-intensity, continuous along the inner

membrane of the nuclear envelope and the cellular membrane, and diffuse low-intensity inside the nucleus, with some sporadic concentrations of protein inside the nuclei. However, the lamin stain was not always homogeneous, showing concentrations of protein in some parts of the nuclear envelope.

All of the nuclei/cells that we used to test the algorithm were successfully segmented. We segmented 19 lamin-stained nuclei from cultured cells, 24 integrin-stained cells from cultured cells and 23 integrin-stained cells in tissue.

Examples of the images, which are 2D images selected from the acquired confocal 3D sets (except 2D-conventional images of real tissue sections) are shown in Figs 4–6, along with the results of the flow superimposed on the original images. Figure 4 shows two examples of the segmentation

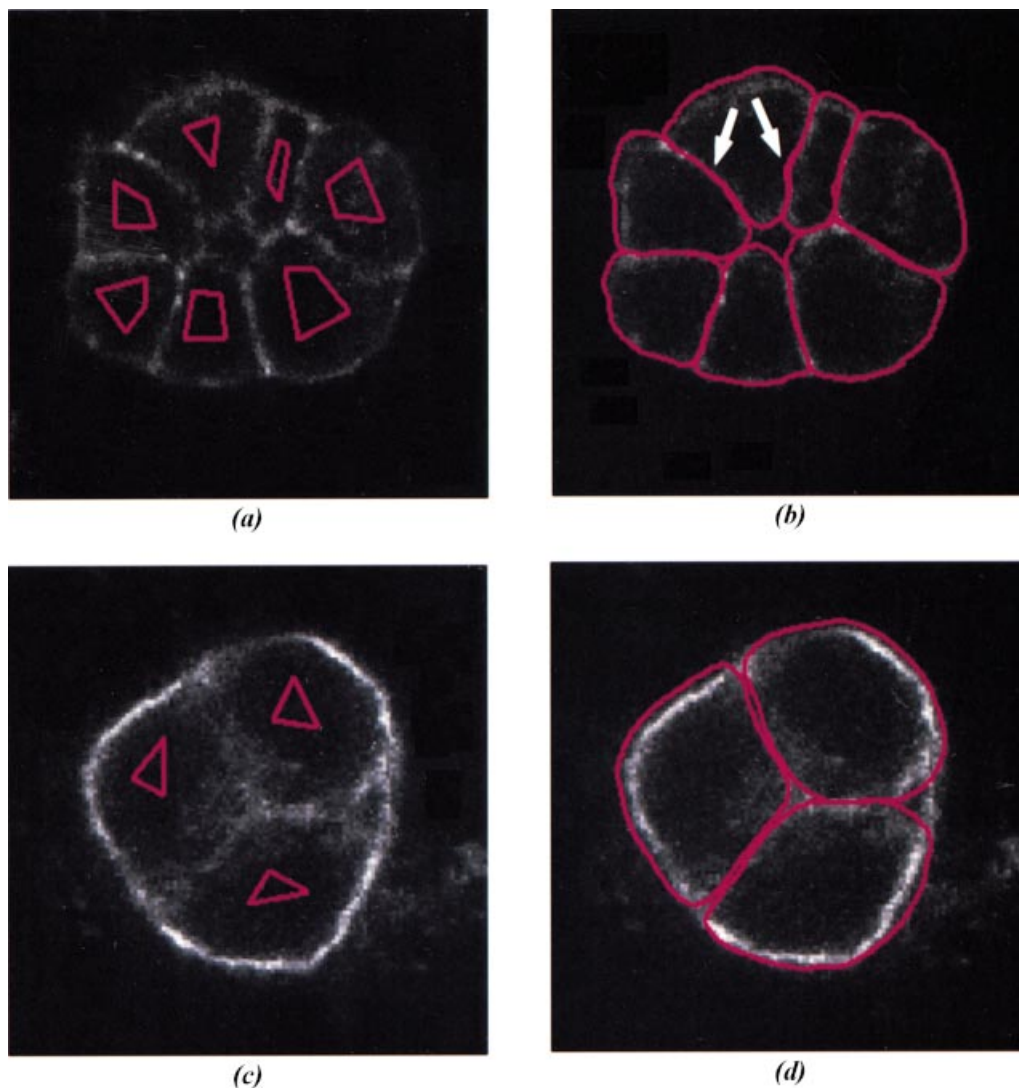


Fig. 5. Examples of the segmentation results on cultured, integrin-stained cells. (a) and (c) show the original images with the interactively drawn seeds. (b) and (d) show the results, superimposed on the original images. The white arrows in (b) show concavities that were followed by the algorithm.

of nuclei in cultured cells stained with lamin antibodies. Figure 5 shows two examples of whole cell segmentation using integrin staining in cultured cells, where the cells formed acini mimicking ducts of the mammary gland. Figure 6 shows whole cell segmentation in intact tissue stained for integrin. Note that both Figs 5 and 6 contain cells with concavities in their surfaces, which the algorithm was able to follow.

These examples demonstrate that the algorithm converges to the nuclear or cellular surface, and that it admits a range of variation in the quality of the staining within and between images. Some tuning of the parameters might be necessary to adapt the segmentation to images substantially different from the images used.

4. Discussion

In this paper we have presented an algorithm for segmenting cells and nuclei based on immunostaining of proteins associated with the plasma and nuclear membranes, respectively. These proteins (lamins for the nuclear envelope and integrins for the cell membrane) precisely delineate the shape of the nuclei and cells. Being able to follow nuclear shape will permit a better understanding of the positions of intranuclear components relative to the nuclear surface in cases where the nuclear shape is contorted. Furthermore, nuclei delineation using lamin or other nuclear-envelope bound proteins is the only segmentation alternative possible when the DNA is condensed or highly heterogeneously

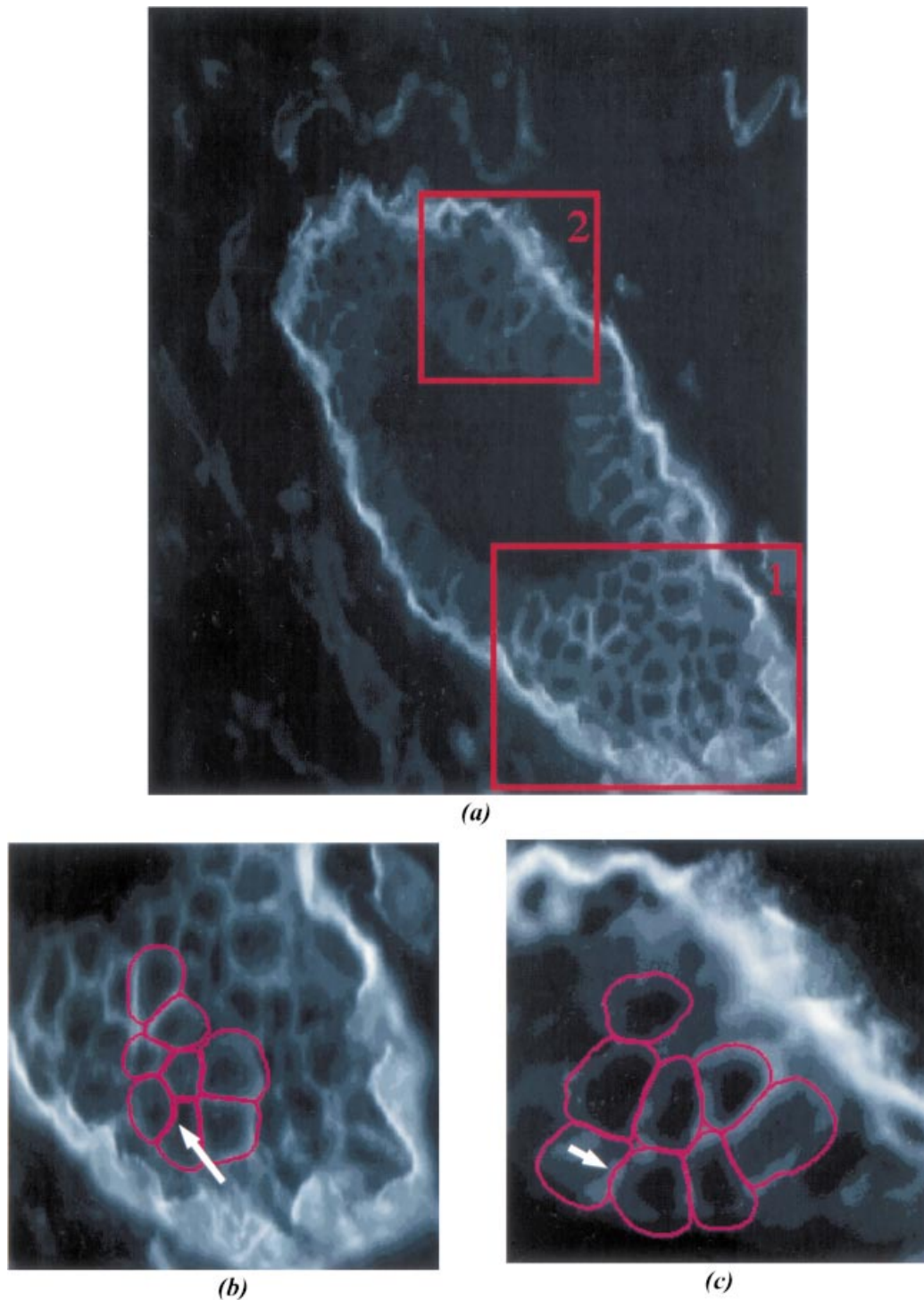


Fig. 6. Example of the segmentation results on two parts (b) and (c) of mouse mammary tissue stained for integrin. Image (a) shows the original image. Images (b) and (c) show the results on some interactively seeded nuclei from two parts of image a. The white arrows in (b) and (c) show concavities that were followed by the algorithm.

distributed within the nucleus. Moreover, lamin staining should permit accurate delineation of the nuclei when the DNA does not completely extend to the nuclear surface, or in those cases where the nuclear surface forms profound

intranuclear invaginations that would be missed using DNA staining alone (Fricker *et al.*, 1997; Clubb & Locke, 1998). Analogous consideration can be made regarding the importance of cell surface delineation.

The selection of these proteins was based on their ubiquity across different types of cells and species and their relatively high concentration in the vicinity of the nuclear envelope and plasma membrane, respectively. In our images, intranuclear lamin staining was sporadically seen, in the form of bright, punctate agglomerations of protein. However, it did not prevent our algorithm from correctly performing. It is also important to emphasize that lamin B is highly flexible in that quality staining can be obtained using several classical fixation methods, and also it attaches to the second (or third) antibody in multicolour assays.

For cell surface segmentation, we chose staining of integrins, even though other markers to proteins or lipids specific to the plasma membrane would be also valid candidates.

Our segmentation method requires the placement of initializing seeds in each nucleus/cell that can be done automatically or interactively. Automatic initialization based on object size estimation works well with objects homogeneous in size and with regular shapes. For more twisted objects, interactive initialization is required, which is kept to a minimum, by just clicking inside each nucleus or cell. However, we have not tested our method on highly irregularly shaped cells such as muscle cells or neurones. When segmenting a binucleate cell, it would be necessary to have only one seed for the cell surface, but two seeds for the two nuclei.

Choosing the values of the parameters for the flows and the number of iterations of the numerical approximation to the solution was done empirically, by tuning them to a model computer-generated image with similar characteristics to the rest of the images in the series that we wanted to analyse. In our experience, the parameters are general enough as to be applicable to specimens prepared in different ways, and with different noise and signal characteristics. In our case, it was also true for different types of staining, since the same values for the parameters used to segment nuclei using lamin staining were also used successfully to segment the integrin-stained cells. An interesting extension of this method would be the development of automatic techniques to tune these parameters for given classes of images.

An obstacle to widespread application of PDEs to image segmentation is that their numerical approximations are computationally expensive, and therefore they are not suitable for real time applications. This problem was partially addressed using narrow band techniques. However, the computation times were still high. Further improvement could be achieved by reducing the domain where each seed is embedded and moves, which in the current implementation is equal to the image size. This would allow reducing the extent of the distance transform associated with each seed, with the consequent memory savings and time optimization. Work on a linearized version of the PDE equation is also in progress which will considerably reduce the

computation time of the algorithm. Once these improvements are implemented, our method could be massively tested and applied for accurate delineation of nuclei or cells prior to quantification of other intranuclear or intracellular components in real images. Extension of these methods to 3D is also being studied. We do not anticipate fundamental problems here, because some of the methods have already been extended for applications to similar images (Ortiz de Solorzano *et al.*, 1999), and the remainder of the methods have been extended to 3D for other applications.

In summary, we have shown this approach to be a good alternative to DNA staining based segmentation of nuclei, in that it provides good segmentation of both isolated and clustered nuclei, and furthermore it was also successful for whole cell segmentation.

Acknowledgements

This work was supported by the Director, Office of Energy Research, Office of Health and Environmental Research of the U.S. Department of Energy under contract NO. DE-AC03-76SF00098 to S.J.L.; the Director Energy's Office of Energy Research, Office of Computational and Technology Research, Mathematical Information, and Computational Sciences Division, Applied Mathematical Sciences Subprogram to R.M.; NIH grant CA-67412 to S.J.L.; NIH grant CA-64786 to Mina J. Bissell; U.S. Department of Defense Breast Cancer Research Program DAMD 17-97-1-7103 to S.A.L. and a contract with Carl Zeiss Inc. We would like to thank Dr Barcellos-Hoff for providing the integrin-stained culture and tissue samples and for her advice regarding surface cell staining, and Dr Mina J. Bissell for providing the S₁₋₅₀ cells.

References

- Alvarez, L., Guichard, E., Lions, P.L. & Morel, J.M. (1993) Axioms and fundamental equations of image processing. *Arch. Ration. Mech. An.* **123**, 197–257.
- Ballard, D.H. (1981) Generalizing the Hough Transform to detect arbitrary shapes. *Pattern Recogn.* **13** (2), 111–122.
- Balzer, P., Furber, A., Cavarro-Menard, C., Croue, A., Tadei, A., Geslin, P., Jallet, P. & Le Jeune, J.J. (1998) Simultaneous and correlated detection of endocardial and epicardial borders on short-axis MR images for the measurement of left ventricular mass. *Radiographics*, **18**, 1009–1018.
- Barcellos-Hoff, M.H. (1993) Radiation-induced transforming growth factor β and subsequent extracellular matrix reorganization in murine mammary gland. *Cancer Res.* **53**, 3880–3886.
- Bosman, F.T. (1993) Integrins: cell adhesives and modulators of cell function. *Histochem. J.* **25**, 469–477.
- Clubb, B.H. & Locke, M. (1998) 3T3 cells have nuclear invaginations containing F-actin. *Tissue Cell*, **30**, 684–691.
- Cohen, L.D. & Cohen, I. (1993) Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. Pattern Anal.* **15**, 1131–1146.

- Dastidar, P., Heinonen, T., Numminen, J., Rautiainen, M. & Laasonen, E. (1999) Semi-automatic segmentation of computed tomographic images in volumetric estimation of nasal airway. *Eur. Arch. Oto-Rhinol.* **256**, 192–198.
- Dean, P., Mascio, L., Ow, D., Sudar, D. & Mullikin, J. (1990) Proposed standard for image cytometry data files. *Cytometry*, **11**, 561–569.
- Dow, A.I., Shafer, S.A., Kirkwood, J.M., Mascari, R.A. & Waggoner, A.S. (1996) Automatic multiparameter fluorescence imaging for determining lymphocyte phenotype and activation status in Melanoma tissue sections. *Cytometry*, **25**, 71–81.
- Elmoataz, A., Schüpp, S., Clouard, R., Herlin, P. & Bloyet, D. (1998) Using active contours and mathematical morphology tools for quantification of immunohistochemical images. *Sig. Proc.* **71**, 215–226.
- Fricker, M., Hollinshead, M., White, N. & Vaux, D. (1997) Interphase nuclei of many mammalian cell types contain deep, dynamic, tubular membrane-bound invaginations of the nuclear envelope. *J. Cell Biol.* **136**, 531–544.
- Irinopoulou, T., Vassy, J., Beil, M., Nicopoulou, Encaoua, D. & Rigaut, J.P. (1997) Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks of prostatic lesions. *Cytometry*, **27**, 99–105.
- Kass, M., Witkin, A. & Terzopoulos, D. (1988) Snakes: active contour models. *Int. J. Comput. Vis.* **1**, 321–331.
- Kikinis, R., Guttman, C.R.G., Metcalf, M.S., Wells, W.M., Gil, J., Ettinger, M.D., Howard, L., Weiner, M.D. & Jolesz, F.A. (1999) Quantitative follow-up of patients with multiple sclerosis using MRI: Technical aspects. *J. Magn. Reson. Imaging*, **9**, 519–530.
- Koukoulis, G.K., Virtanen, I., Korhonen, M., Laitinen, L., Quarana, V. & Gould, V.E. (1991) Immunohistochemical localization of integrins in the normal, hyperplastic and neoplastic breast. *Am. J. Pathol.* **139**, 787–799.
- Lelièvre, S.A., Weaver, V.M., Nickerson, J.A., Larabell, C.A., Bhaumik, A., Peterson, O.W. & Bissell, M.J. (1998) Tissue phenotype depends on reciprocal interactions between the extracellular matrix and the structural organization of the nucleus. *Proc. Natl. Acad. Sci.* **95**, 14711–14716.
- Lockett, S.J. & Herman, B. (1994) Automatic detection of clustered, fluorescence-stained nuclei by digital image-based cytometry. *Cytometry*, **17**, 1–12.
- Lockett, S.J., Sudar, D., Thompson, C.T., Pinkel, D. & Gray, J.W. (1998) Efficient, interactive, three-dimensional segmentation of cell nuclei in thick tissue sections. *Cytometry*, **31**, 275–286.
- Malladi, R. & Sethian, J. (1996) Image processing: Flows under Min/Max curvature and mean curvature. *Graph Models Image Processing*, **58**, 127–141.
- Malladi, R. & Sethian, J. (1997) Level set methods for curvature flow, image enhancement, and shape recovery in medical images. *Visualization and Mathematics: Experiments, Simulations and Environments* (ed. by H C. Hege and K. Polthier), pp. 329–345. Springer Verlag, Heidelberg.
- Malladi, R. & Sethian, J. (1998) A real-time algorithm for medical shape recovery. *Proceedings ICCV'98, Mumbai India, January*, pp. 304–310.
- Malladi, R., Sethian, J. & Vemuri, B.C. (1995) Shape modeling with front propagation: a level set approach. *IEEE Trans. Pattern Anal.* **17**, 158–175.
- Malpica, N., Ortiz de Solorzano, C., Vaquero, J.J., Santos, A., Vallcorba, I., Garcia-Sagredo, J.M. & del Pozo, F. (1997) Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*, **28**, 289–297.
- Mikula, K., Sarti, A. & Lamberti, C. (1997) Geometrical diffusion in 3D echocardiography. *Proceedings of ALGORITMY'97 Conference on Scientific Computing, West Tatra Mountains, Slovakia*.
- Ortiz de Solorzano, C., Garcia Rodriguez, E., Jones, A., Pinkel, D., Gray, J.W., Sudar, D. & Lockett, S.J. (1999) Segmentation of confocal microscope images of cell nuclei in thick tissue section. *J. Microsc.* **193**, 212–226.
- Osher, S.J. & Sethian, J.A. (1988) Fronts propagation with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. P.* **79**, 12–49.
- Petersen, O.W., Ronnov-Jessen, L., Howlett, A.R. & Bissell, M.J. (1992) Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. *Proc. Natl. Acad. Sci. USA*, **89**, 9064–9068.
- Rigaut, J.P., Vassy, J., Herlin, P., Duigou, F., Masson, E., Briane, D., Foucrier, J., Carvajal-Gonzalez, S., Downs, A.M. & Mandard, A.M. (1991) Three-dimensional DNA image cytometry by confocal scanning laser microscopy in thick tissue blocks. *Cytometry*, **12**, 511–524.
- Rodenacker, K., Aubele, M., Hutzler, P. & Umesh Adiga, P.S. (1997) Groping for quantitative digital 3-D image analysis: an approach to quantitative in situ hybridization in thick tissue sections of prostate carcinoma. *Anal. Cell. Pathol.* **15**, 19–29.
- Romeny, B.M., ed. (1994) *Geometry-Driven Diffusion in Computer Vision*. Kluwer, Dordrecht.
- Sarti, A., Mikula, K. & Sgallari, F. (1999) Nonlinear multiscale analysis of 3D echocardiographic sequences. *IEEE Trans. Med. Imaging*, **18**, 453–466.
- Sethian, J. (1997) *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Sciences*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge.
- Sochen, N., Kimmel, R. & Malladi, R. (1998) A general framework for low level vision. *IEEE Trans. Image Process., Special Issue on PDEs and Geometry-Driven Diffusion in Image Processing and Analysis*, **7**, 310–318.
- Vincent, L. (1993) Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE Trans. Image Process.* **2**, 176–201.